

Agreement Between AI-Detected Coronary Artery Calcium & Expert Visual Assessment: A Multi-Site Analysis

Camila V. Blair, MD^a · Daniel M. Huck, MD MPH^a · Milena Petranovic, MD^b · Arthur Shiyovich, MD^a · Rhanderson Cardoso, MD^a · Stephanie A. Besser, MS^a · Olivia Gabriel, BA^a · Brittany Weber, MD PhD^c · Joanne Miao, BS^a · Yotam Kimmel, BS^{c,d} · Orit Wimpfheimer, MD^d · Kelley R. Branch, MD^e · Michael D. Shapiro, DO^f · Ron Blankstein, MD^a

^aBrigham and Women's Hospital, Harvard Medical School · ^bMassachusetts General Hospital · ^cUT Southwestern · ^dNanox Imaging · ^eUniversity of Washington · ^fWake Forest University School of Medicine



Mass General Brigham



Introduction

Opportunistic **CAC detection on routine chest CT** is a scalable strategy for cardiovascular risk stratification. However, the reliability of **automated CAC detection** across heterogeneous clinical settings remains uncertain.

AI algorithm
automated scoring

Expert reader
visual assessment

Hypothesis / Objectives

Evaluate **agreement between an FDA-cleared AI CAC algorithm and expert visual assessment** across three clinical sites, using clinically relevant CAC thresholds.

Methods

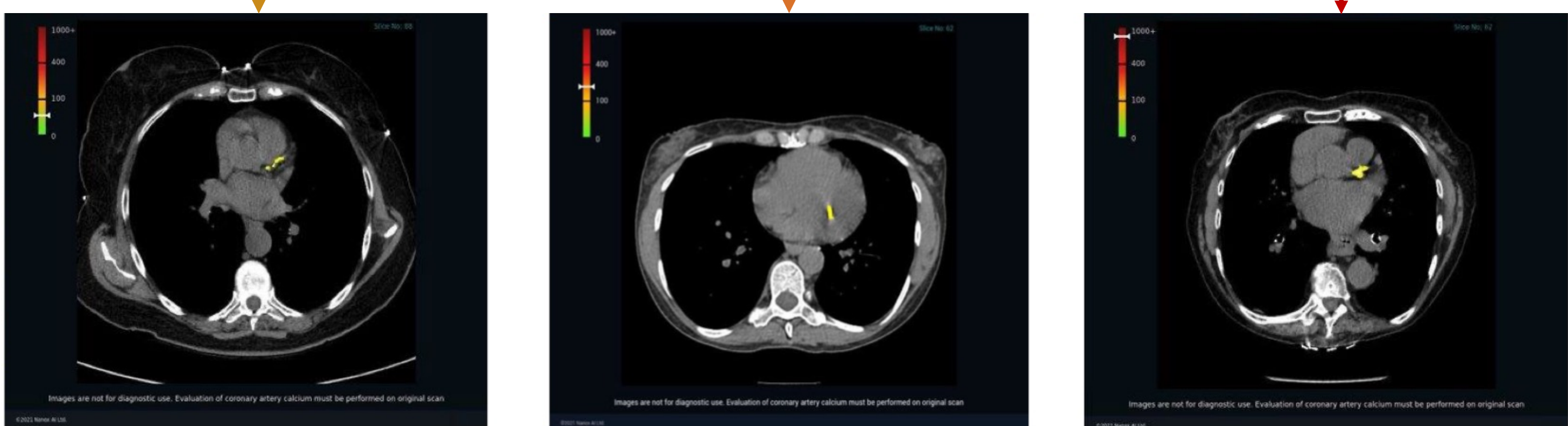
NON-CONTRAST CHEST CT · 3 INSTITUTIONS

AI
HealthCCSng 2.0
(Nanox.AI), FDA-cleared

Expert
Site-level visual
assessment

CAC SEVERITY CATEGORIES (AGATSTON)

ZERO 0 LOW 1–99 MED 100–399 HIGH ≥400



CHEST CT
STUDIES
5,468
across 3 clinical sites

- Agreement & disagreement scored per CAC category
- Subgroup: AI CAC = 0 vs visual (1,290 studies, 2 sites)
- Excluded: studies with metal artifacts

Conclusion

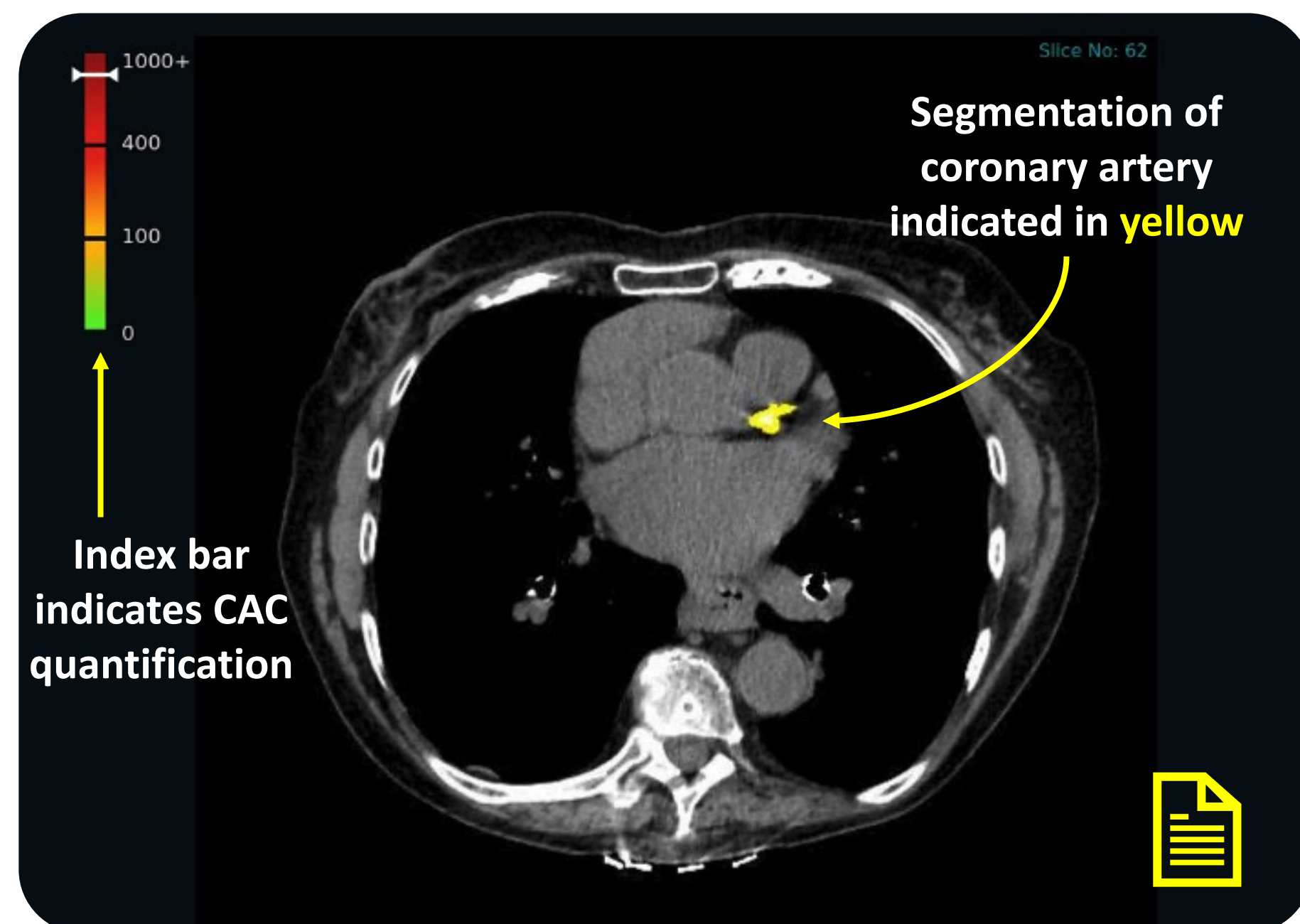
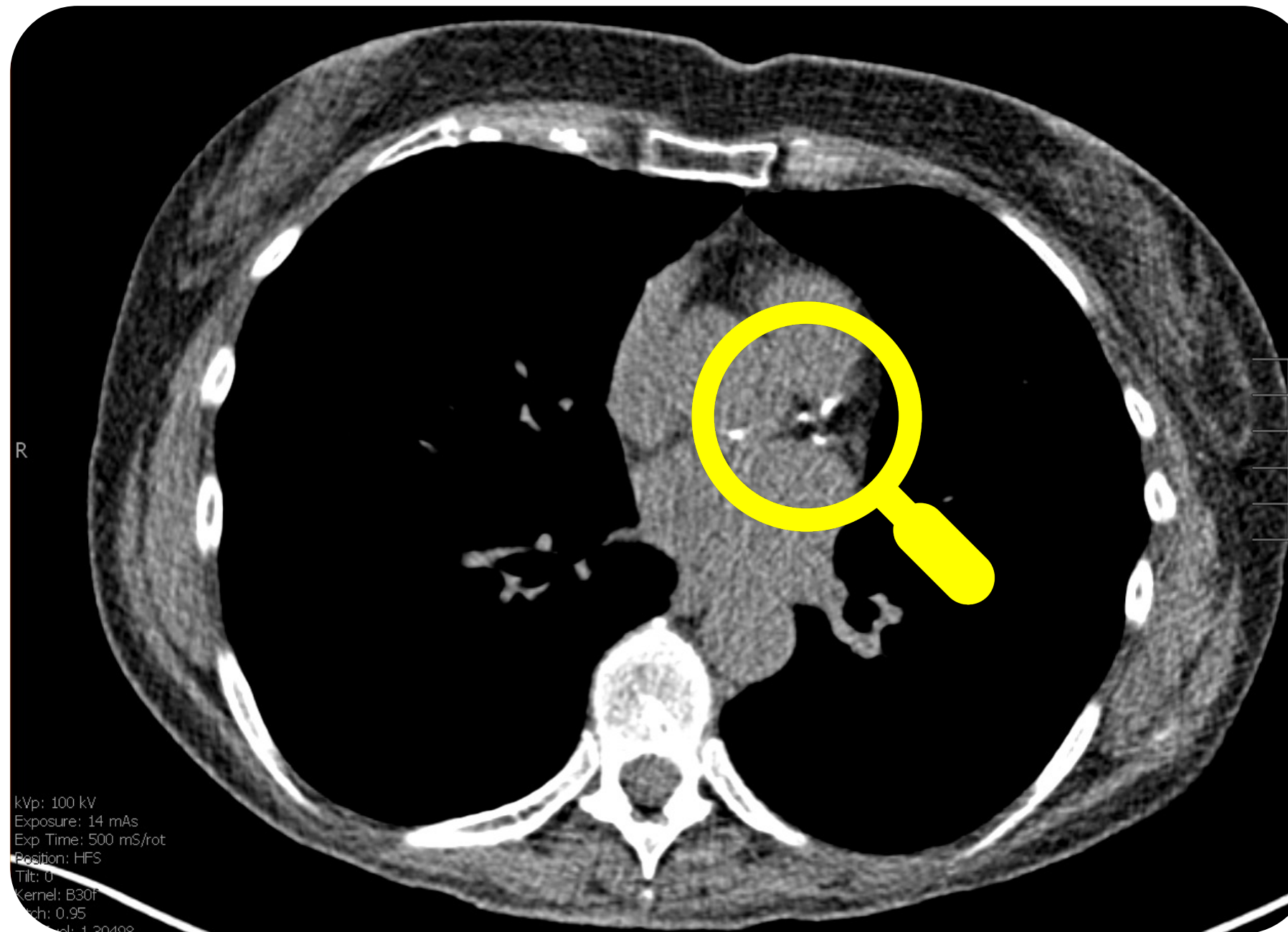
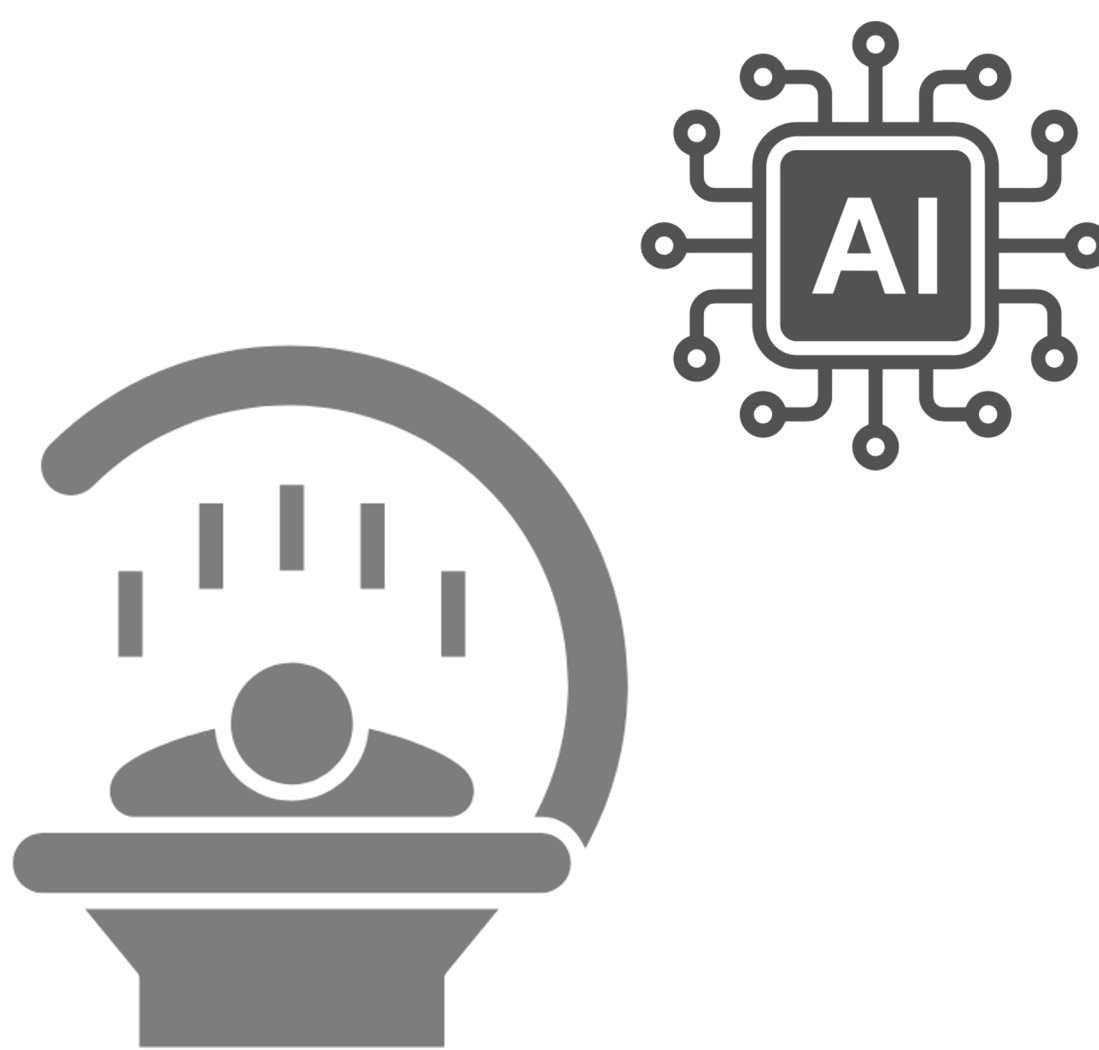
AI-based detection of incidental CAC shows **high agreement with expert visual assessment across three clinical sites**. Future studies are warranted to evaluate its impact on patient management and outcomes.

Results

>90%
overall agreement, all CAC categories

3
independent clinical sites

5,468
non-contrast chest CT studies



AI–expert agreement by site & CAC category

	Site 1	Site 2	Site 3
CAC = 0	99.0%	—	96.2%
CAC > 0 · any	90.0%	90.1%	90.7%
Low · 1–99	92.9%	88.6%	85.5%
Medium · 100–399	87.4%	89.7%	96.1%
High · ≥400	85.7%	93.3%	96.9%

Table 1 · cell = % AI–expert agreement; Site 2 CAC = 0 not assessed

CLINICALLY RELEVANT THRESHOLDS

CAC ≥ 100 91.4%
CAC ≥ 400 92.6%

Agreement at the thresholds that drive statin & prevention decisions

CAC = 0 SUBGROUP

98.0%
agreement on AI-classified CAC = 0

1,290
studies · 2 sites

26
discordant cases

WHERE THEY DISAGREE

7.1–14.5%

Disagreement was most common in the **low-CAC** category most often due to **mitral annular calcification**.

DISCLOSURES

CAC quantified by **HealthCCSng 2.0 (Nanox.AI)**, FDA-cleared. Y. Kimmel & O. Wimpfheimer are employees of Nanox Imaging.

CONTACT

cvblair@bwh.harvard.edu
rblankstein@bwh.harvard.edu